working paper
department
of economics

AN IMPROVED RATE FOR NONNEGATIVE DEFINITE
CONSISTENT COVARIANCE MATRIX ESTIMATION WITH
HETEROGENEOUS DEPENDENT DATA

Danny Quah

No. 529 July 1989

massachusetts
institute of
technology
50 memorial drive
cambridge, mass.02139

AN IMPROVED RATE FOR NONNEGATIVE DEFINITE CONSISTENT COVARIANCE MATRIX ESTIMATION WITH HETEROGENEOUS DEPENDENT DATA

Danny Quah

No. 529 July 1989

An Improved Rate for Nonnegative Definite Consistent
Covariance Matrix Estimation With Heterogeneous Dependent Data

by

Danny Quah *
July 1989.

* Department of Economics, MIT and NBER (*dquah@dolphin.mit.edu*). I thank Jeffrey Wooldridge for discussions, and the MIT Statistics Center for its hospitality. Wooldridge also kindly pointed out a serious mistake in a first draft. All errors and misinterpretations are mine.

An Improved Rate for Nonnegative Definite Consistent Covariance Matrix Estimation With Heterogeneous Dependent Data

by
Danny Quah
Economics Department, MIT.
July 1989.

Abstract

This paper improves on previous rates at which lag lengths are allowed to grow for consistent covariance matrix estimation with heterogeneous dependent data. Using a WLLN, we give a consistency result for growth rates of $o(n^{1/3})$*; the previous rate was* $o(n^{1/4})$*. This new rate equals that of Berk's autoregressive spectral density estimator for well-behaved stationary contexts, and thus may be best possible outside of very special cases.*

1. Introduction

Estimating consistent covariance matrices is one of the most common problems confronting the applied researcher. The need to do this arises in econometric work ranging from Euler equation estimation by GMM methods (Hansen and Singleton [1982]) to tests for integration and cointegration (Phillips [1987], Phillips and Perron [1988], Phillips and Ouliaris [1988], and Stock [1988]). Thus the results of Hansen [1982] (for stationary data), and White [1984], White and Domowitz [1984], and Newey and West [1987] (for heterogeneous dependent data) on consistent covariance matrix estimation have been very widely applied.

Newey and West [1987] adapted results in White [1984] and White and Domowitz [1984] to obtain a class of non-negative definite consistent covariance matrix estimators for dependent non-iid data. Their correction to arguments in White [1984] 6.19 led to a $o(n^{1/4})$ rate for increasing lag length to preserve consistency. The same rate appears in Gallant and White [1988] 6.18, and has been used quite generally (see for instance Phillips [1987] and Phillips and Perron [1988]).

This paper improves that rate to $o(n^{1/3})$, the same as that for Berk's autoregressive (time domain) spectral density estimator for strictly stationary data. Except for very special cases, this new rate of $o(n^{1/3})$ may be the best possible. Since the choice of lag length is often one of the most troubling (and seemingly arbitrary) to applied researchers, this rate improvement should permit greater flexibility without risking the

loss of consistent inference.

Note that this $o(n^{1/3})$ rate is exactly that originally in the conclusion of White's [1984] Theorem 6.20. As Newey and West [1987] have pointed out however, this result did not follow from White's proof. This paper therefore uses a different argument to re-establish that result, under essentially the same regularity assumptions. The proof is remarkably straightforward.

2. Notation

The p-norm of a random variable (rv) X defined on a given probability space $(\Omega, \mathcal{F}, \Pr)$ is denoted $||X||_p \stackrel{\text{def}}{=} E^{1/p}|X|^p$. The random variable that is the absolute value of X is denoted $|X|$. Recall that $\{X_t,\ t \geq 1\}$ is said to be *α-mixing of size $-q$* if the α-mixing coefficients tend to zero and satisfy $\alpha_m = O(m^\lambda)$ for some $\lambda < -q$, so that $\sum_m \alpha_m^{1/q} < \infty$. Similarly X is said to be *ϕ-mixing of size $-q$* if its ϕ-mixing coefficients satisfy analogous conditions. See for example Gallant and White [1988]. Let ϕ^R denote the ϕ-mixing coefficients when X_t is reversed in time; let $\phi_m^+ = \max(\phi_m, \phi_m^R)$. When X is Gaussian and covariance stationary, $\phi_m^+ = \phi_m = \phi_m^R$ for all m. It will be convenient below to place restrictions on ϕ^+.

We will need the following:

Lemma 2.1 (Davydov's Inequality): *For all p, q such that $\frac{1}{p} + \frac{1}{q} < 1$,*

$$|EX_tX_{t-j} - EX_tEX_{t-j}| \leq 15\alpha_j^{1-\frac{1}{p}-\frac{1}{q}}||X_t||_p \cdot ||X_{t-j}||_q.$$

∎

See for example Philipp [1986, p.241] Lemma 3.1 for this form of Davydov's result.

Lemma 2.2 (Peligrad's Inequality): *For all p, q such that $\frac{1}{p} + \frac{1}{q} = 1$,*

$$|EX_tX_{t-j} - EX_tEX_{t-j}| \leq 2(\phi_j)^{1/p}(\phi_j^R)^{1/q}\ ||X_t||_p\ ||X_{t-j}||_q.$$

∎

This was first obtained in Peligrad [1983], and improves by $(\phi_j^R)^{1/q}$ on the earlier long-standing inequality. White [1984] 6.16 is a special case of 2.1; when X is covariance stationary, 2.2 is a strict improvement on White [1984] 6.16.

3. Results

The first result is a weak law of large numbers (*WLLN*) for a process that fails the usual weak dependence assumptions for mixingales (and thus for mixing sequences as well). Further the process will have growing first absolute moments so that it is not an L^1-mixingale (Andrews [1988]).

We give the regularity assumptions in two sets, one set of assumptions on the process itself, and the other on a set of weights.

Assumption 3.1: *Suppose* $\{X_t,\ t \geq 1\}$ on $(\Omega, \mathcal{F}, \Pr)$ *satisfies* $EX_t = 0$ *for all* t, *and assume further that for some* $r > 1$: *(i.)* $\sup_t \|X_t\|_{4r} < \infty$; *and (ii.) either* (a.) X_t *is* α*-mixing of size* $-2r/(r-1)$ or *(b.)* X_t *is* ϕ^+*-mixing of size* -2. ∎

For convenience in notation, let $X_t = 0$ for all $t \leq 0$.

Assumption 3.2: *Suppose* $w_n(j)$, *with* $n \geq 1$, $j \geq 0$ *is a double array of uniformly bounded non-negative weights such that as* $n \to \infty$, *we have* $w_n(j) \to 1$ *for each* j. ∎

These assumptions are essentially those in Newey and West [1987] Theorem 2, or White [1984] Theorem 6.20 where applicable.

The first result is a *WLLN* for dependent double arrays that will be used below.

Theorem 3.3: *Assume (3.1) and (3.2), and define the double array of rv's:*

$$Z_{nt} \stackrel{\text{def}}{=} \sum_{j=0}^{l(n)} w_n(j)\,(X_t X_{t-j} - EX_t X_{t-j})$$

for some sequence of nonnegative integers $l(n)$, *with* $l(n) = o(n^{1/3})$. *Then the double array* $\{Z_{nt}\}$ *satisfies* $n^{-1}\sum_{t=1}^{n} Z_{nt} \xrightarrow{P} 0$ as $n \to \infty$. ∎

Remarks

1. Notice that if $l(n) \uparrow \infty$, Z_{nt} has stronger long term dependence than does a mixingale. Further, Z_{nt} will have growing moments for $l(n)$ increasing with n.
2. Clearly the conclusion remains true if $l(n)$ is fixed, or if Z_{nt} is defined to exclude the $j = 0$ term.
3. Our improved rate derives from using this *WLLN* below in place of the implication rule as in White's [1984] proof of his Theorem 6.20.

4. By first giving this *WLLN*, it should be clear that our proof differs from that in White [1984] Chapter 6, by a change in the order of summation.

The principal result is convergence in probability for a weighted estimator of $\text{Var}\left(n^{-1/2}\sum_{t=1}^{n} X_t\right)$. We state this as follows:

Theorem 3.4: *Assume (3.1) and (3.2), and let $l(n)$ be a sequence of positive integers such that as $n \to \infty$, $l(n) \uparrow \infty$, and $l(n) = o(n^{1/3})$. Then as $n \to \infty$,*

$$n^{-1}\left[\sum_{t=1}^{n} X_t^2 + 2\sum_{j=1}^{l(n)} w_n(j) \sum_{t=j+1}^{n} X_t X_{t-j}\right] - \text{Var}\left(n^{-1/2}\sum_{t=1}^{n} X_t\right) \xrightarrow{p} 0.$$

∎

Remarks

1. Apply Theorem 3.4 to the proof of Theorem 2 in Newey and West [1987] to argue convergence of the second and third terms in their expression (9). The other terms similarly converge to zero by $l(n) \uparrow \infty$ and $l(n) = o(n^{1/3})$. The result here therefore implies the same conclusion as their Theorem 2, but with greater flexibility in choice of lag length ($o(n^{1/3})$ instead of $o(n^{1/4})$).
2. Similarly, the results in Chapter 6 of Gallant and White [1988] remain intact with their *Assumption TL* changed to $m_n = o(n^{1/3})$ from the typographically incorrect $O(n^{1/4})$ on p.101.
3. The Newey-West result is widely used in applications: see for example Phillips [1987] Theorem 4.2, Phillips and Perron [1987], and Phillips and Ouliaris [1988], and elsewhere. Those results therefore all hold with an even more flexible choice for the lag length.
4. The rate $o(n^{1/3})$ is also that used in autoregressive spectral density estimation under assumptions on X that include strict stationarity, absolute summability of the Wold moving average coefficients, and finite fourth moments on the iid innovations (e.g. Berk [1974] Theorem 1).
5. Fuller [1976, Theorem 7.2.3] and Anderson [1971, Chapter 9] imply that a $o(n)$ rate can be used for the strictly stationary case. It may in fact be possible to adapt the "unraveling" method used there for the nonstationary mixing situation considered here.

4. Proofs

In the sequel, the symbols K and K' will denote arbitrary finite constants, not necessarily the same throughout.

Our first result is a *WLLN*. It is convenient to give the proof in two parts, the first part is a variance bound which may be useful in other applications. This bound is also that part of the results that gives the binding restriction on the mixing and moment conditions (3.1); thus if further improvement is forthcoming, it is likely to obtain by giving a sharper inequality here.

Proof of Theorem 3.3: First bound $\mathrm{Var}\left(\sum_{t=1}^{n} Z_{nt}\right)$. Write $\mathrm{Var}\left(\sum_{t=1}^{n} Z_{nt}\right) = \left|\sum_{t=1}^{n}\sum_{s=1}^{n} EZ_{nt}Z_{ns}\right|$. Decompose this double sum of products into products close together, and products far apart. By the triangle inequality,

$$\mathrm{Var}\left(\sum_{t=1}^{n} Z_{nt}\right) \leq \sum_{t} \sum_{|s-t|\leq 2l(n)} |EZ_{nt}Z_{ns}| + \sum_{t} \sum_{|s-t|>2l(n)} |EZ_{nt}Z_{ns}|.$$

Consider the first summand. For s, t between 1 and n, the Cauchy-Schwarz inequality implies:

$$|EZ_{nt}Z_{ns}| \leq ||Z_{nt}||_2 \cdot ||Z_{ns}||_2 \leq \sup_{1\leq t\leq n} ||Z_{nt}||_2^2.$$

For each t, there are at most $4l(n)+1$ points s for which $|s-t| < 2l(n)$. Thus the first summand satisfies:

$$\sum_{t} \sum_{|s-t|\leq 2l(n)} |EZ_{nt}Z_{ns}| \leq n \cdot (4l(n)+1) \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_2^2.$$

Next consider the second summand. For $|s-t| > 2l(n)$, Davydov's Inequality implies:

$$|EZ_{nt}Z_{ns}| \leq 15\alpha_{l(n)}^{1-\frac{1}{r}} \cdot ||Z_{nt}||_{2r} \cdot ||Z_{ns}||_{2r} \leq 15\alpha_{l(n)}^{1-\frac{1}{r}} \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_{2r}^2,$$

while Peligrad's Inequality implies:

$$|EZ_{nt}Z_{ns}| \leq 2\phi_{l(n)}^{+} \cdot ||Z_{nt}||_2 \cdot ||Z_{ns}||_2 \leq 2\phi_{l(n)}^{+} \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_2^2.$$

There are at most $n-4l(n)-1$ points s such that $|s-t| > 2l(n)$. Thus the second summand obeys:

$$\sum_{t} \sum_{|s-t|>2l(n)} |EZ_{nt}Z_{ns}| \leq n(n-4l(n)-1) \cdot 15\alpha_{l(n)}^{1-\frac{1}{r}} \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_{2r}^2 \leq n^2 \cdot 15\alpha_{l(n)}^{1-\frac{1}{r}} \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_{2r}^2.$$

Similarly, the second summand also satisfies:

$$\sum_t \sum_{|s-t|>2l(n)} |EZ_{nt}Z_{ns}| \leq n^2 \cdot 2\phi^+_{l(n)} \cdot \sup_{1\leq t\leq n} ||Z_{nt}||_2^2.$$

For any p such that $2 \leq p \leq 2r$, Minkowski's inequality gives $||Z_{nt}||_p \leq \sum_{j=0}^{l(n)} w_n(j)||X_tX_{t-j} - EX_tX_{t-j}||_p$. By 3.1.i, $\sup_t \sup_j ||X_tX_{t-j} - EX_tX_{t-j}||_p \leq \infty$. Further, since $w_n(j)$ is uniformly bounded, we have that for some finite constant K, $||Z_{nt}||_p \leq K \cdot (l(n)+1) \Rightarrow ||Z_{nt}||_p^2 \leq K^2 \cdot (l(n)+1)^2$. Using this for $p = 2$ and $2r$ in the first and second summands, conclude that there must exist some finite constant K' such that:

$$\text{Var}\left(\sum_{t=1}^n Z_{nt}\right) \leq K'\left\{n \cdot (4l(n)+1) \cdot (l(n)+1)^2 + n^2(l(n)+1)^2 \cdot \alpha_{l(n)}^{1-\frac{1}{r}}\right\}$$

and

$$\text{Var}\left(\sum_{t=1}^n Z_{nt}\right) \leq K'\left\{n \cdot (4l(n)+1) \cdot (l(n)+1)^2 + n^2(l(n)+1)^2 \cdot \phi^+_{l(n)}\right\}.$$

If 3.1.ii.a, then $\alpha_j = O\left(j^\lambda\right)$ for some $\lambda < -2r/(r-1)$, so that $\alpha_{l(n)}^{1-\frac{1}{r}} = O\left(l(n)^{\lambda'}\right)$ for some $\lambda' < -2$, or $l(n)^2\alpha_{l(n)}^{1-\frac{1}{r}} = o(1)$. Similarly, if 3.1.ii.b, then $l(n)^2\phi^+_{l(n)} = o(1)$. But then, using Chebyshev's inequality, for any $\epsilon > 0$,

$$\begin{aligned}\Pr\left[n^{-1}|\sum_{t=1}^n Z_{nt}| \geq \epsilon\right] &\leq \frac{K'}{\epsilon^2}n^{-2} \cdot \text{Var}\left(\sum_{t=1}^n Z_{nt}\right)\\ &\leq \frac{K'}{\epsilon^2}\left\{n^{-1} \cdot (4l(n)+1) \cdot (l(n)+1)^2 + (l(n)+1)^2 \cdot \alpha_{l(n)}^{1-\frac{1}{r}}\right\}\end{aligned}$$

and

$$\Pr\left[n^{-1}|\sum_{t=1}^n Z_{nt}| \geq \epsilon\right] \leq \frac{K'}{\epsilon^2}\left\{n^{-1} \cdot (4l(n)+1) \cdot (l(n)+1)^2 + (l(n)+1)^2 \cdot \phi^+_{l(n)}\right\}.$$

Given $l(n) = o(n^{\frac{1}{3}})$, and 3.1.ii, one or the other of the right hand sides above tend to zero as $n \to \infty$. Therefore, as $n \to \infty$, $n^{-1}\sum_{t=1}^n Z_{nt} \xrightarrow{p} 0$. *Q.E.D.*

Next, turn to the main result (3.4). The proof of this is an abbreviation and modification of ideas in Newey and West [1987] and White [1984] Chapter 6. The crucial difference is in replacing White's (corrected) Lemma 6.19, the implication rule, and an early use of Chebyshev's Inequality with our *WLLN* for double arrays (3.3). With this *WLLN*, (3.4) follows in a remarkably straightforward way.

Proof of Theorem 3.4: Proceed in two steps. First, argue the expected version with truncation and weighting differs from $\text{Var}(n^{-1/2}\sum_{t=1}^n X_t)$ by a quantity that vanishes as $n \to \infty$. Then show the feasible

estimator converges in probability to its expectation. Begin with the expected version:

$$n^{-1}\left[\sum_{t=1}^{n} EX_t^2 + 2\sum_{j=1}^{l(n)} w_n(j) \sum_{t=j+1}^{n} EX_tX_{t-j}\right] - \mathrm{Var}\left(n^{-1/2}\sum_{t=1}^{n} X_t\right)$$

$$= \frac{2}{n}\sum_{j=1}^{l(n)}(w_n(j)-1)\sum_{t=j+1}^{n} EX_tX_{t-j} - \frac{2}{n}\sum_{j=l(n)+1}^{n-1}\sum_{t=j+1}^{n} EX_tX_{t-j}.$$

By Davydov's Inequality,

$$|EX_tX_{t-j}| \le 15\alpha_j^{1-\frac{1}{2r}}\|X_t\|_{4r}\cdot\|X_{t-j}\|_{4r},$$

and by Peligrad's Inequality,

$$|EX_tX_{t-j}| \le 2\phi_{l(n)}^{+}\|X_t\|_2\cdot\|X_{t-j}\|_2.$$

By 3.1, $\sup_t \|X_t\|_{4r} < \infty$ and $\sup_t \|X_t\|_2 < \infty$, so that:

$$|n^{-1}\sum_{j=1}^{l(n)}(w_n(j)-1)\sum_{t=j+1}^{n} EX_tX_{t-j}| \le K\sum_{j=1}^{l(n)}|\left(\frac{n-j}{n}\right)(w_n(j)-1)|\alpha_j^{1-\frac{1}{2r}} \le K\sum_{j=1}^{l(n)}|w_n(j)-1|\alpha_j^{1-\frac{1}{2r}},$$

and similarly,

$$|n^{-1}\sum_{j=1}^{l(n)}(w_n(j)-1)\sum_{t=j+1}^{n} EX_tX_{t-j}| \le K\sum_{j=1}^{l(n)}|w_n(j)-1|\phi_{l(n)}^{+}.$$

From 3.1.ii, we have that either $\sum_{j=1}^{\infty}\alpha_j^{\frac{1}{2}-\frac{1}{2r}} < \infty$ or $\sum_{j=1}^{\infty}(\phi_{l(n)}^{+})^{\frac{1}{2}} < \infty$ which imply either $\sum_{j=1}^{\infty}\alpha_j^{1-\frac{1}{2r}} < \infty$ or $\sum_{j=1}^{\infty}\phi_{l(n)}^{+} < \infty$, respectively. Since $w_n(j) \to 1$ for each j, the dominated convergence theorem then implies that $\sum_{j=1}^{l(n)}|w_n(j)-1|\alpha_j^{1-\frac{1}{2r}} \to 0$ or $\sum_{j=1}^{l(n)}|w_n(j)-1|\phi_{l(n)}^{+} \to 0$ as $n \to \infty$. By a similar argument, we have that:

$$\left|n^{-1}\sum_{j=l(n)+1}^{n-1}\sum_{t=j+1}^{n} EX_tX_{t-j}\right| \le K\sum_{j=l(n)+1}^{n-1}\left(\frac{n-j}{n}\right)\alpha_j^{1-\frac{1}{2r}} \le K\sum_{j=l(n)+1}^{n-1}\alpha_j^{1-\frac{1}{2r}},$$

and

$$\left|n^{-1}\sum_{j=l(n)+1}^{n-1}\sum_{t=j+1}^{n} EX_tX_{t-j}\right| \le K\sum_{j=l(n)+1}^{n-1}\phi_{l(n)}^{+}.$$

Again, if 3.1.ii.a, $\sum_{j=1}^{\infty}\alpha_j^{1-\frac{1}{2r}}$ converges to a finite quantity. Similarly, if 3.1.ii.b, $\sum_{j=1}^{\infty}\phi_{l(n)}^{+}$ converges to a finite quantity. In either case, this implies the respective right hand side above converges to 0, provided that $l(n)\uparrow\infty$ as $n\to\infty$. This completes the first part of the proof.

Second, we show the required convergence in probability. The estimator

$$n^{-1}\left[\sum_{t=1}^{n} X_t^2 + 2\sum_{j=1}^{l(n)} w_n(j)\sum_{t=j+1}^{n} X_tX_{t-j}\right]$$

differs from its expectation by:

$$n^{-1}\left[\sum_{t=1}^{n}(X_t^2 - EX_t^2) + 2\sum_{j=1}^{l(n)} w_n(j) \sum_{t=j+1}^{n} (X_tX_{t-j} - EX_tX_{t-j})\right].$$

For ease of notation, define $X_t = 0$ for all $t \leq 0$, and rearrange orders of summation in the second term:

$$2n^{-1}\sum_{j=1}^{l(n)} w_n(j) \sum_{t=j+1}^{n} (X_tX_{t-j} - EX_tX_{t-j}) = 2n^{-1}\sum_{t=1}^{n}\sum_{j=1}^{l(n)} w_n(j)(X_tX_{t-j} - EX_tX_{t-j}).$$

Define the double array of rv's $Z_{nt} = \sum_{j=1}^{l(n)} w_n(j)(X_tX_{t-j} - EX_tX_{t-j})$, and apply Theorem 3.3 to it. The term $n^{-1}2\sum_{j=1}^{l(n)} w_n(j)\sum_{t=j+1}^{n}(X_tX_{t-j} - EX_tX_{t-j})$ therefore converges in probability to zero. Similarly apply Theorem 3.3 (with $l(n) = 0$) to $X_t^2 - EX_t^2$, so that the first term $n^{-1}\sum_{t=1}^{n}(X_t^2 - EX_t^2)$ converges in probability to zero as well. This completes the proof. *Q.E.D.*

Notice that the first part of the proof uses considerably weaker mixing-moment conditions than necessary for the second part of the proof, which is essentially a repeated application of Theorem 3.3. Thus an improvement in the lag growth rate, if forthcoming, might most easily be found by obtaining a sharper inequality than that available in the proof of Theorem 3.3.

References

Andrews, Donald W.K. (1988): "Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables," *Econometric Theory*, 4, no. 3, December, 458-467.

Anderson, T.W. (1971): *The Statistical Analysis of Time Series*, New York: John Wiley and Sons.

Berk, K.N. (1974): "Consistent Autoregressive Spectral Estimates," *Annals of Statistics*, 2 no.3, 489-502.

Fuller, W.A. (1976): *Introduction to Statistical Time Series*, New York: John Wiley and Sons.

Gallant, A.R. and H. White (1988): *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, New York: Basil Blackwell.

Hansen, L.P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, no.4, July, 1029-1054.

Hansen, L.P. and K.J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, no. 5, September, 1269-1286.

McLeish, D.L. (1975): "A Maximal Inequality and Dependent Strong Laws," *Annals of Probability*, 3, no.5, 829-839.

Newey, W.K. and K.D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, no.3, May, 703-708.

Peligrad, M. (1983): "A Note on Two Measures of Dependence and Mixing Sequences," *Adv. Appl. Prob.*, 15, 461-464.

Philipp, W. (1986): "Invariance Principles for Independent and Weakly Dependent Random Variables," pp. 225-268 in E. Eberlein and M.S. Taqqu (eds.), *Dependence in Probability and Statistics*, Boston: Birkhauser.

Phillips, P.C.B. (1987): "Time Series Regression with A Unit Root," *Econometrica*, 55, no.2, March, 277-301.

Phillips, P.C.B. and S. Ouliaris (1986): "Testing for Cointegration," Cowles Foundation Discussion Paper no. 890, Yale University.

Phillips, P.C.B. and P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 78, no.2, 335-346.

Stock, J.H (1988): "A Class of Tests for Integration and Cointegration," Kennedy School of Government, Harvard University mimeo, March.

White, H. (1984): *Asymptotic Theory for Econometricians*, New York: Academic Press.

White, H., and I. Domowitz (1984): "Nonlinear Regression with Dependent Observations," *Econometrica*, 52, 143-161.